2007

What to Ask in a Self-Administered Dietary Assessment Website: The Role of Professional Judgement

Y. C. Probst
University of Wollongong, yasmine@uow.edu.au

Linda C. Tapsell
University of Wollongong - Dubai Campus, ltapsell@uow.edu.au

What to Ask in a Self-Administered Dietary Assessment Website: The Role of Professional Judgement

Abstract

Background: Professional judgement is a key component of questionnaire development, subjective in nature and rarely reported in this context. It is required in dietary questionnaires to delimit the size whilst retaining quality of the data.
Objectives: To describe the nature and extent of professional judgement involved in developing a food database to include in a web-based self-administered dietary assessment.
Methods & materials: Professional judgement was applied in tandem with a stepwise statistical analysis of hierarchically reported foods in the Australian National Nutrition Survey (NNS95). Statistical analyses determined foods commonly consumed and eaten together and three different forms of cluster analysis were then used to group foods that were most similar in macronutrient content. Professional judgement was required to interpret these groupings and determine the most suitable clustering technique. Face validity of the resulting food groups was determined by recognition of the food name by experienced dietitians, as usually reported in a diet history interview.
Results: Applying professional judgement to differentiate between foods after the cluster analysis resulted in an increase from 370 to 501 food groups. A final three-level hierarchy of 19, 103 and 422 groups in the new database compared with 21, 106 and 370 groups of NNS95 was developed.
Conclusions: The use of professional judgement in database development is an important step when they are to be used in self-adminstered assessments. It ensures foods are not only nutritionally appropriate but also conceptually appropriate for recognition by a layperson.

Keywords

Dietary assessment, Questionnaire development, Internet, Statistical analysis, Professional judgement

Disciplines

Arts and Humanities | Life Sciences | Medicine and Health Sciences | Social and Behavioral Sciences

Publication Details

This article was originally published as Probst, YC & Tapsell, L, What to Ask in a Self-Administered Dietary Assessment Website: The Role of Professional Judgement, Journal of Food Composition and Analysis, 20(8), 2007, 696-703.

What to Ask in a Self-Administered Dietary Assessment Website: The Role of Professional Judgement

Yasmine Probst*

Smart Foods Centre, University of Wollongong, Northfields Avenue, Wollongong, Australia.

Phone: +612 4221 5302, Fax: +612 4221 4844, Email: yasmine@uow.edu.au

Linda Tapsell

National Centre of Excellence in Functional Foods, University of Wollongong, Northfields Avenue, Wollongong, Australia.

Phone: +612 4221 3152, Fax: +612 4221 4844, Email: ltapsell@uow.edu.au

* Corresponding author

Abstract

Background: Professional judgement is a key component of questionnaire development, subjective in nature and rarely reported in this context. It is required in dietary questionnaires to delimit the size whilst retaining quality of the data.

Objectives: To describe the nature and extent of professional judgement involved in developing a food database to include in a web-based self-administered dietary assessment.

Methods & materials: Professional judgement was applied in tandem with a stepwise statistical analysis of hierarchically reported foods in the Australian National Nutrition Survey (NNS95). Statistical analyses determined foods commonly consumed and eaten together and three different forms of cluster analysis were then used to group foods that were most similar in macronutrient content. Professional judgement was required to interpret these groupings and determine the most suitable clustering technique. Face validity of the resulting food groups was determined by recognition of the food name by experienced dietitians, as usually reported in a diet history interview.

Results: Applying professional judgement to differentiate between foods after the cluster analysis resulted in an increase from 370 to 501 food groups. A final three-level hierarchy of 19, 103 and 422 groups in the new database compared with 21, 106 and 370 groups of NNS95 was developed.

Conclusions: The use of professional judgement in database development is an important step when they are to be used in self-adminstered assessments. It ensures foods are not only nutritionally appropriate but also conceptually appropriate for recognition by a layperson.

Keywords:

Dietary assessment

Questionnaire development

Internet

Statistical analysis

Professional judgement

Introduction

The development of a database for nutrient analysis is a complex process involving vast amounts of data (Ireland et al., 2002). In Australia, for example, such databases can contain over 4500 individual food items (each with their own item codes) for which nutrient data is available (Australia New Zealand Food Authority, 1999). The methodology employed to organise this data into a useful format will also vary depending on the type of output required and the overall use of the database. Although the use of statistical analysis is a primary methodology in the development of food databases (Shai et al., 2004, Peterson and Dwyer, 2000, Akbaya et al., 2000). many studies do not report the methods used in the database development (Ireland et al., 2002). Studies that do report the analysis vary widely. For example Akbaya et al. utilised hierarchical cluster analysis to develop a composition database of lamb. The cluster analysis helped to determined the differences in the fat composition of lamb prepared using different methods (Akbaya et al., 2000).

Food composition databases however are vital to dietary assessment methodology. An automated diet history interview was developed in the Illwarra region of New South Wales, Australia allowing individuals with metabolic syndrome to self-report their usual dietary intake. The most recently reported data on Australian dietary intakes is provided in the National Nutrition Survey (NNS95) of 1995. This survey provides 24-hour recall data for 13858 individuals. In the study reported here, foods collected from the NNS95 were used, sorted into a database of four-level hierarchical food groups. Each level of the hierarchy varied in the level of detail about the food items with the broadest level containing 21 food groupings and the most detailed level containing more than 4500 individual food items. This format would allow for a multiple-pass format and allow users To log out of the assessment and return at a later time. In this project the NNS95 data underwent various phases of statistical analysis, between which professional interpretation of the results were required.

In this context, professional judgement, the application of knowledge, skills, values and experiences of qualified professionals to the interpretation of data, has not been widely reported. A Medline (New York, N.Y.: Ovid Technologies, Columbia) search from 1966 to 2005 found that studies utilising professional judgment as a methodology either do not define the details or have been reported more than ten years ago (Bentsen et al., 1988, Farand et al., 1995, Gilmore, 1992, Hepworth, 1989, Regan, 1981, Slavkin, 1972). The few recent studies describing such processes (Greaves and Grant, 2000, Coles, 2002, Lelie et al., 1999, Lo et al., 2005) do not commonly relate to the field of nutrition, and none to database development. The only nutrition paper found, described a comparison of professional judgement used by dietitians and dietetic technicians with an algorithm used for assessment of malnutrition in hospital patients. Finding differences between the level of experience and the reliability of the professional judgement, the inter-rater reliability of the algorithm was preferred (Lowery et al., 1998).

Therefore, the aim of this paper is to describe the nature and extent of dietetic professional judgement involved in developing a food database for inclusion in the user interface of the self-administered dietary assessment website.

Methods

Raw food and nutrient data from the NNS95, sorted by meals, was initially used to determine the degree of error associated with the use of individual food names representing groups of foods. The analysis revealed that grouping foods resulted in a 5-10% reporting error. Resultantly a total of 370 foods listed in the second NNS95 level were selected as the starting point for analysis and database development rather than the 4500+ individual food items.

A list of foods commonly consumed per meal, was established by determining those consumed by 99% of the population for both frequency of consumption and contribution to total energy. Foods eaten together (associated foods) were also assessed based on a 50% confidence level. The entire food list then underwent cluster analysis to group foods based on similarities in macronutrient composition.

Output data from statistical analysis of food intake data reported in the NNS95 (Burden et al.) was interpreted by an experienced dietitian upon completion of each analysis. The following outlines the professional judgement that was required. The interpretation of the data was based on recognition of foods by the dietitian as food commonly reported during a diet history interview (due to the age of the nutrition survey data), or readily available to the general public in major retail outlets. Foods were included if they were identifiable in documentation from previously conducted diet history interviews from intervention studies {Martin, 2003 #3325; Tapsell, 2004 #8519} using a similar target population. Foods were excluded if they were unavailable to the general public or if they were aimed at population groups outside of the target group, for example children or infants. Further, new foods were added to the database if the dietitian considered them to be consumed in greater quantities today than ten years ago based on the intervention study data. Food eaten together and the composition of the food groups based on nutrient and conceptual similarities were then considered. This was important as not only did the user need to recognise the food items but the groupings need to be useful to the dietitian who would receive the output data from the website. Finally the devised food database was assessed by 5 dietitians for face validity and modified by consensus.

Common and associated foods

The item codes of the NNS95 food groups were used in the statistical analysis (Burden et al.). Names of the food groups were not included in the analysis; therefore, interpretation of the results began with naming of each of the item codes that emerged from the statistical analysis for each meal (breakfast, lunch, dinner and snacks). Each common food was then related back to the original food grouping from NNS95. The analysis resulted in a number of individual food names remaining that were not identified as commonly eaten e.g. plain flour. This list of foods was then reviewed to determine the impact of their inclusion or exclusion within the food database for the study population (i.e. adults with metabolic syndrome).

The names of associated foods were to be used as probing questions in the web-based questionnaire via a questioning hierarchy. For example if a user selects cereal for breakfast, the website would then automatically ask for foods that are eaten with the cereal such as milk or fruit. While food groups had been previously alphabetised in the NNS95, the associated food lists obtained from further statistical analysis of the NNS95 data, needed to be review by the dietitian to determine their inclusion or exclusion within the question schedule for the automated dietary assessment. Food combinations were included if they were reported together in diet history interviews. Foods which did not appear in the output for associated foods statistical analysis but appeared regularly in the diet history documents from intervention studies were also added to the food list. Foods that were not commonly reported together were excluded from the associated foods listing, yet were still included in the common foods listing if they met the inclusion criteria for that category.

Re-grouping the foods

The cluster analysis was used to regroup the food items based on macronutrient similaries and the professional judgement was applied to ensure that these groupings were also conceptually similar. Cluster analysis data was originally provided for the entire list of more than 4500 foods (Burden et al.). Interpretation of the statistical data revealed unrelated foods to be grouped together ie. High in carbohydrate but not conceptually similar, and clustering needed to be selective. Therefore individual groups from the NNS95 database were clustered separately. These groups were chosen by assessment of the original NNS95 food groups for nutrient and conceptual. Foods which were judged as not recognisable by a layperson were re-clustered individually.

Output from the cluster analysis took the form of dendrogram plots displaying the stepwise progression of groups from the least similar as a whole (leaf nodes) to those most similar as a whole (stems). Dependant on group size, tables were also provided using Microsoft Excel (version 2000, Microsoft Corporation, USA). These tables indicated the split of the food groups when dendrograms were indistinguishable (Burden et al.).

A cut-off point for the number of groups to be formed, needed to be created for each dendrogram produced. This point needed to ensure that a 'picture' of the entire food group could be seen and the majority of foods which were similar were all within the cluster. The cut-off was determined by listing each of the leaf nodes (individual food items) of the dendrogram in order of their appearance and assessing similarities in nutrient composition for carbohydrate, protein and fat (saturated fat, monounsaturated fat and polyunsaturated fat). Foods that appeared both nutritionally and conceptually similar at the lowest level of the dendrogram were grouped based on the grouping node to which it relates. For example Figure 1 provides dendrogram output for group 127 (breakfast cereals) based on the average linkage clustering technique. The food code numbers (leaf nodes) were each related back to their corresponding food name shown in the table. The grouping arms for the foods were then followed up on the dendrogram stem until the grouping of foods were considered recognisable to the general public. These foods were then linked back to the original groupings used by NNS95. If these groups already existed, they were added to the new food database. If a group of these food items did not already exist, a new group was created using a generic name of all foods contained within the group. A similar process was followed for the tables of clustered groups for which the dendrogram output was unclear. Each statistical output, whether dendrogram or tabular gave three different forms of clustering, through the Ward, average linkage and complete linkage methods respectively. Interpretation required each method to be assessed separately and the method where, the grouping of foods that were not only similar in nutrient composition but conceptually similar and relevant to the layperson, was selected. Often this required combinations of more than one technique to be used. This resulted in a subjective grouping of foods based on their positioning in the cluster.

Face validity analysis

An assessment of face validity of the composition and names of the final food groups was then conducted. As the data from the NNS95 survey was prepared for researchers and statisticians, the naming criteria of the food groups did not necessarily reflect those used by laypersons to describe foods. Similarly the generic names used from the cluster analysis were not considered suited to the layperson.

A total of five experienced dietitians from three states of Australia (New South Wales, Victoria and Queensland) reviewed the final food database to ensure that food names were those referred to by the layperson, and that all commonly reported foods, from their experience of the diet history interview, were included using the existing food database from the statistical analysis. The two dietitians from Victoria renamed food items and expanded/condensed food groups to warrant ease of finding the foods by the end-user. The two dietitians from Queensland simultaneously created their own list of food items and grouped them accordingly. Upon completion of the two separate food lists, all dietitians met and critically analysed the resultant lists. Many food groups were combined to minimise the large number of food groups available. Group consensus also resulted in foods that were not usually eaten alone being removed from the food database (eg. Flour). These foods were either moved to the associated foods list or eliminated from the database.

The final two food lists were combined by the New South Wales dietitian. Foods from the second list were assessed for their inclusion in the existing food database. The primary decision process related to the level of grouping of a food item. Many of the foods from the second food list were individual food items and needed to be grouped based on nutrient or conceptual similarities or were able to be inserted under one of the existing groups in the food database. Once all items from the second food list had been considered, the resulting food groups in the database again needed to be renamed for recognition by the layperson. This renaming involved the addition of example food items recognisable by the general public.

Results

Common and associated foods

Each of the food items of the NNS95 analysis were found to be compatible with those considered commonly eaten at mealtimes by the study population. Of the 4551 individual foods of NNS95 only 3519 were reported during the survey (Burden et al.). Using these foods, milk was found as the most commonly reported food item by frequency of reporting and also by contribution to total energy (Table 1). [INSERT TABLE 1]

Infant formulas and infant foods were excluded from the database as they were not suited to the study population decreasing the total number of individual food items to 3500.

Due to changes in the food supply over the past decade, the option of 'other' was included in the database for each NNS95 food group. This would allow newly developed foods to be categorised accordingly with time.

Similarly the associated food list identified foods that are currently eaten together today. Milk with breakfast cereal appeared 24% of the time for all milk associations (n=111), tea with sugar appeared 23% of the time for all tea associations (n=53) and bread with margarine appeared 24% of the time for all bread associations (n=54). Due to the number of foods that were associated with more than one food item, the process of ordering the food database based on associated foods was eliminated. Each meal was given a standard order of foods which could be modified once the foods were in the website. Due to the overlap between the numbers of associated foods, a list of foods that were eaten with another food were compiled (Table 2) and were to be used as prompting questions in the corresponding location of the website. [INSERT TABLE 2]

Regrouping the foods

The original cluster analysis of the entire 4500 plus foods resulted in, for example, noodles and custard being grouped together due to their carbohydrate similarities (Burden et al.). These were not believed to be conceptually similar. Of the 21 upper level groups of NNS95 10 groups were re-clustered. Within each of these upper level groups, groups from the second level of NNS95 were used for the clustering. For example within the milk products and dishes category of NNS95, cheeses were re-clustered separately from milks and yoghurts due to the conceptual similarities. Table 3 shows the results from the cluster analysis and the final interpretation for the food database. The sample shown in the table signifies the importance of using more than one clustering technique. The groups provided by a single cluster analysis on its own were not suitable to a database for use by the layperson. Using a combination of each analysis resulted in groups which were not only nutritionally but also conceptually similar. [INSERT TABLE 3]

Applying professional judgement to the food database resulted in an increase from the original 370 third level NNS95 food groups to 453 food groups. This change in numbers was primarily the result of separating groups into their fatty acid constituents, the clustering output separating foods such as toasted bread or bread rolls from the untoasted forms and the addition of ‘other’ categories.

Face validity analysis

The food list created by the Queensland dietitians primarily contained individual food items which could be related back to the original 4500 plus item food list as a cross reference for inclusion within the database. For example, Herbs, spices, flours, custard and baking powder and gelatine were eliminated as they were not eaten alone. This decreased the list of individual food items 3437 foods to be grouped.

Following a standardised format, food groups in the database needed to be expanded further to include forms currently available in the marketplace and not seen in the NNS95 data of 1995. The breads and bread rolls group for example contained subcategories for white, wholemeal, mixed grain and rye breads, whereas the English muffins group after clustering did not. This group was therefore modified to include the newly available wholemeal and mixed grain muffins. This further increase in food group numbers resulted in 501 groups in total.

Ninety-two percent of foods from the NNS95 database were renamed, for example 'Breakfast cereal, biscuit, regular, whole wheat, low sugars' became 'Wheat based biscuits eg Weet bix, Vita Weets'. Foods were primarily renamed to simplify the description and to add an example food to allow recognition of the group by the layperson.

Table 4 shows the change of the food groups from the original NNS95 food groups, to those used in this study. It can be seen that these upper level groupings result in less groups due to the reallocation of some groups to the associated foods list, while others were excluded altogether. The final food group numbers were different from those of NNS95 (Table 5), primarily due to differences in the timing of the database creation and also the primary function of our database compared with that of NNS95.

[INSERT TABLE 4][INSERT TABLE 5]

Discussion

The creation of a food database for patients to self-report their food intake saw many challenges including consistency of grouping and level of detail of the final food database. Although there has been a significant change in the types and brands of food items available to the general public (Williams et al., 2003), the generic groupings of food items has not changed significantly. This was identified in the present study through interpretation of both common food items of a 1995 national survey and the sorting of associated food items from this same survey. When comparing the final groups from the study with those assessed by Ireland et al (2002) it may be seen that the food groups created in each of the studies assessed (including this one) follow a common trend in food group composition, with the naming of the groups primarily differing dependant on the end-user. Furthermore, it should be highlighted that the individual food items and nutrient data of the groups did need to be modified based on the population and country within which the database will be used. The food groups from this study appear most similar with those of the Euro Food Groups (EFG), though the groups used in Europe were mono-hierarchical seeing foods from both our database and associated foods list included within the hierarchy (Ireland et al., 2002).

If only the statistical output from the NNS95 had been used, the end-users' understanding of questions would have been limited. Consideration needed to be given to ensure that the primary focus of the re-arrangement of the NNS95 food groups was not purely based on nutrient data that would be understood by a dietitian. The underlying assumption that held throughout this study was that the average end-user would not be highly food literate and therefore the database would also needed to consider the types of wording chosen and the amount of detail of the individual food groups. This was limited by the level of experience of the dietitians and their own subjective interpretation of patients' understanding. Inclusion of brand name food products as examples was therefore used in an attempt to overcome misinterpretation by the user.

The extent of professional judgement required for the creation of a food database for self-administered dietary assessment was primarily seen in the form of face validity testing to ensure the food names could be easily identified and understood by the user. Such data has not previously been reported, though informatics literature commonly focuses on the use of simplified terminology in user interface design (Tanriverdi and Jacob, 2001, Suàrez et al., 2004, Hartson and Hix, 1989). Description of the nature of professional judgement used in the creation of the food database was found to be a highly complex process as similarly found by Greaves and Grant (2000). Despite this complexity, the inclusion of professional judgement for food database design should be encouraged in the literature to assist those developing similar applications.

The resultant food database will undergo further modification when it is uploaded to the dietary assessment website, as the screen size and layout will also need to be considered. The database will then be tested with the end-user under both laboratory and clinical settings. Outcomes of these further studies will help to determine the validity of the professional judgement described in this study. The use of automated assessment processes is only one of the many expanding areas of technology within dietetics. Already utilised in psychology (Crespin and Austin, 2002)and dentistry (Abbey et al., 2003) such technology can be used to assist rather than supplement practice, though without clear processes for developing food databases we cannot ensure the end-user benefits will be maximised.

REFERENCES

ABBEY, L. M., ARNOLD, P., HALUNKO, L., HUNEKE, M. B. & LEE, S. (2003) CASE STUDIES for Dentistry: development of a tool to author interactive, multimedia, computer-based patient simulations. *Journal of Dental Education,* 67**,** 1345-54.

AKBAYA, A., ELHANB, A., ÖZCANC, C. & DEMIRTASD, S. (2000) Hierarchical cluster analysis as an approach for systematic grouping of diet constituents on basis of fatty acid, energy and cholesterol content: application on consumable lamb products. *Medical Hypotheses,* 55**,** 147-154.

AUSTRALIA NEW ZEALAND FOOD AUTHORITY (1999) *AUSNUT - Australian Food and Nutrient Database,* Canberra, ANZFA.

BENTSEN, B. S., MICHALSEN, H. & FOLLERAS, S. (1988) Social-medical aspects of cystic fibrosis in Norway. IV. A comparison of the parents' and the professionals' judgement of the severity of the handicap. *Scandinavian Journal of Gastroenterology - Supplement,* 143**,** 65-7.

BURDEN, S., PROBST, Y., STEEL, D. & TAPSELL, L. (Under Review) Identification of food descriptors for use in theoretical development of a computer-assisted diet history interview. *Journal of the American Dietetic Association*.

COLES, C. (2002) Developing professional judgment. *Journal of Continuing Education in the Health Professions,* 22**,** 3-10.

CRESPIN, T. R. & AUSTIN, J. T. (2002) Computer technology applications in industrial and organizational psychology. *Cyberpsychology & Behavior,* 5**,** 279-303.

FARAND, L., LEPROHON, J., KALINA, M., CHAMPAGNE, F., CONTANDRIOPOULOS, A. P. & PREKER, A. (1995) The role of protocols and professional judgement in emergency medical dispatching. *European Journal of Emergency Medicine,* 2**,** 136-48.

GILMORE, A. (1992) Ottawa's meningococcal outbreak provided a lesson in professional judgement and science. *CMAJ Canadian Medical Association Journal,* 147**,** 729-32.

GREAVES, J. & GRANT, J. (2000) Watching anaesthetists work: using the professional judgement of consultants to assess the developing clinical competence of trainees. *British Journal of Anaesthesia,* 84**,** 525-533.

HARTSON, H. R. & HIX, D. (1989) Human-computer interface development: concepts and systems for its management. *ACM Computing Surveys (CSUR),* 21**,** 5 - 92.

HEPWORTH, S. (1989) Professional judgement and nurse education. *Nurse Education Today,* 9**,** 408-12.

IRELAND, J., VAN ERP-BAART, A. M. J., CHARRONDIERE, U. R., MOLLER, A., SMITHERS, G. & TRICHOPOULOS, A. (2002) Selection of a food classification system and a food composition database for future food consumption surveys. *European Journal Clinical Nutrition,* 56**,** S33-S54.

LELIE, A., DEKKERS, W., VAN HAMERSVELT, H., HUYSMANS, F. & TEN HAVE, H. (1999) Essay. Discontinuing dialysis: patient's wishes and professional judgement. *Nephrology Dialysis Transplantation,* 14**,** 318-321.

LO, B., DORNBRAND, L. & DUBLER, N. (2005) HIPAA and patient care: the role for professional judgment. *Journal of the American Medical Association,* 293**,** 1766-1771.

LOWERY, J., HILLER, L., DAVIS, J. & SHORE, C. (1998) Comparison of professional judgment versus an algorithm for nutrition status classification. *Medical Care,* 36**,** 1578-1588.

PETERSON, J. & DWYER, J. (2000) An Informatics Approach to Flavonoid Database Development. *Journal of Food Composition and Analysis,* 13**,** 441-454.

REGAN, W. A. (1981) Emergency nursing and professional judgement [case report]. *Regan Report on Nursing Law,* 21**,** 1.

SHAI, I., SHAHAR, D., VARDI, H. & DRORA, F. (2004) Selection of food items for inclusion in a newly developed food-frequency questionnaire. *Public Health Nutrition,* 7**,** 745-749.

SLAVKIN, H. C. (1972) Professional judgement a "must" in nutritional awareness. *San Fernando Valley Dental Society Bulletin,* 6**,** 8-9 passim.

SUÀREZ, P. R., JÙNIOR, B. L. & DE BARROS, M. A. (2004) Applying knowledge management in UI design process. *Proceedings of the 3rd annual conference on Task models and diagrams.* Prague, Czech Republic, ACM Press, New York, USA.

TANRIVERDI, V. & JACOB, R. J. K. (2001) VRID: a design model and methodology for developing virtual reality interfaces. *Proceedings of the ACM symposium on Virtual reality software and technology table of contents.* Baniff, Alberta, Canada, ACM Press, New York, USA.

WILLIAMS, P., MCMAHON, A. & BOUSTEAD, R. (2003) A case study of sodium reduction in breakfast cereals and the impact of the Pick the Tick food information program in Australia. *Health Promotion International,* 18**,** 51-56.

Table 1: Foods commonly reported in NNS95 showing percentage of all food items reported.

Meal	*Frequency of consumption*	*%*	*Contribution to total energy*	*%*
Breakfast	Milk	11.1	Milk	13.0
Lunch	Tomatoes	3.8	White bread	6.7
Dinner	Carrots	3.1	Rice	3.2
Snacks	Milk	8.4	Milk	5.7

Table 2: List of associated food groups used in prompting questions in the web-based questionnaire.

Bread	Margarine	Salad filling
Butter	Mayonnaise	Sauces & gravies
Cheese	Meat filling	Savoury spread e.g. Vegemite
Cream	Milk	Sour cream
Dip	Oil	Sugar & sweetener
Egg filling	Pasta	Sweet sauce & topping
Fish filling	Potatoes	Sweet spread e.g. Jam
Fruit	Rice	Syrup
Malt extract e.g. Milo	Salad dressing	Yoghurt

Table 3: Sample of cluster analysis for NNS95 group 194 (Cheese), showing macronutrients and results for each separate cluster technique (Burden et al.) and areas of professional judgement. Numbers bolded for the clustering technique indicate group used for formation of final food group.

NNS95 Food item	***Macronutrient Composition (g)***				***a Clustering technique***			***Final food group***
	Energy (kJ)	***Carb***	***Pro***	***Fat***	***bWard***	***c Avg***	***dComp***	
Cheese, cream, reduced fat	803	3.1	8.4	16.5	**3**	2	2	Cream cheese, cream cheese based-dips, fruit cheeses
Dip, cream cheese-based, reduced fat, commercial	682	12.9	4.6	10.8	**3**	2	2	Cream cheese, cream cheese based-dips, fruit cheeses
Cheese, bocconcini	856	0.1	17.2	15.2	4	2	**1**	Other soft cheeses
Cheese, goat	823	1.0	13.1	15.8	4	2	**1**	Other soft cheeses
Cheese, haloumi	1020	1.8	21.3	17.1	4	2	**1**	Other soft cheeses
Cheese, processed, cheddar type, reduced fat (fat > 12%)	1066	7.0	17.7	17.6	4	2	**1**	Other soft cheeses

Cheese spread, cheddar-based, reduced fat	995	6.5	16.5	16.5	4	2	**1**	Other soft cheeses
Cheese, mozzarella	1260	0.1	26.0	22.0	4	**1 ***	2	Full fat cheese eg. Cheddar, parmesan, mozzarella
Cheese, pizza	1300	0.1	28.8	21.9	4	**1 ***	2	Full fat cheese eg. Cheddar, parmesan, mozzarella
Cheese, cheddar, reduced fat (25% reduction)	1370	0.0	28.7	23.8	4	**1**	2	Reduced fat cheese eg. Light cheese, 25% reduced fat
Cheese, edam, reduced fat	1290	0.1	33.0	19.8	4	**1**	2	Reduced fat cheese eg. Light cheese, 25% reduced fat
Cheese, gouda, reduced fat	1354	0.1	30.8	22.4	4	**1**	2	Reduced fat cheese eg. Light cheese, 25% reduced fat
Cheese, mozzarella, reduced fat	1200	0.1	31.7	17.9	4	**1**	2	Reduced fat cheese eg. Light cheese, 25% reduced fat
Cheese, Swiss, reduced fat	1390	0.1	34.7	21.6	4	**1**	2	Reduced fat cheese eg. Light

								cheese, 25% reduced fat
Cheese, reduced fat, NFS	1346	0.0	29.0	23.1	4	**1**	2	Reduced fat cheese eg. Light cheese, 25% reduced fat
Cheese, cheddar, reduced fat (50% reduction)	1107	0.0	31.3	15.5	4	**2**	2	Reduced fat cheese eg. Light cheese, 25% reduced fat
Cheese, cheddar, low fat	844	0.1	33.9	7.2	4	**2**	2	Reduced fat cheese eg. Light cheese, 25% reduced fat
Cheese, feta, reduced fat	974	0.1	25.7	14.5	4	**2**	2	Reduced fat cheese eg. Light cheese, 25% reduced fat
Cheese, processed, cheddar type, reduced fat (fat < 12%)	829	3.7	24.0	9.8	4	**2**	2	Reduced fat cheese eg. Light cheese, 25% reduced fat
Cheese, fat-modified, reduced cholesterol	1394	0.1	34.0	22.0	4	**3**	3	Soy cheese, Lo Chol, Mini Chol

* Professional judgement required

a Numbers shown correspond with the group number formed i.e. all foods under one clustering technique with the same number were determined as similar by the clustering technique to which it corresponds

b The Ward method uses sum of square to minimise the distance between any two clusters to create exclusive subsets which are internally similar with respect to the specified criteria. (Ward 1963)

c Average distance clustering technique calculates the distance between clusters are determined by the average distance between any two subsets (Stockburger 2001)

d Complete linkage clustering technique calculates the distances between clusters are determined by the furthest distance between any two subsets (Stockburger 2001)

Abbreviations: NNS95 – National Nutrition Survey, Carb – Carbohydrate, Pro – Protein, NFS – No form specified, Ward – Ward method, Avg – Average Linkage method, Comp – Complete Linkage method.

Table 4: First level food groups showing original NNS95 food groups from which they were formed.

1st level NNS95 food groups	*'New' 1st level food groups*
Non alcoholic beverages	Non-alcoholic drinks
Cereals and cereal products	Rice & pasta dishes
	Bread
	Cereal
Cereal-Based Products and Dishes	Biscuits and crackers
	Convenience and takeaway foods
	Bakery products
Fats and Oils	*
Fish and Seafood Products and Dishes	Dishes with meat, chicken or fish
Fruit Products and Dishes	Fruit & fruit dishes
Egg Products and Dishes	Eggs & egg dishes
Meat, Poultry & Game Products and Dishes	Dishes with meat, chicken or fish
	Meat, chicken & fish (not in a dish)
Milk Products and Dishes	Dairy
Soup	Soups
Seed and Nut Products and Dishes	Savoury snack foods
Savoury Sauces and Condiments	
Vegetable Products and Dishes	Vegetables and vegetable dishes
	Salad
Legume and Pulse Products and Dishes	Vegetables and vegetable dishes
Snack Foods	Savoury snack foods

Sugar Products and Dishes	Sweet snack foods
Confectionary and Health Bars	Savoury snack foods
Alcoholic Beverages	Alcoholic drinks
Special Dietary Foods	Meal replacements & supplements
Miscellaneous	*
Infant Formulae and Foods	**

* Moved to associated food groups, ** Excluded

Abbreviations: NNS95 – National Nutrition Survey

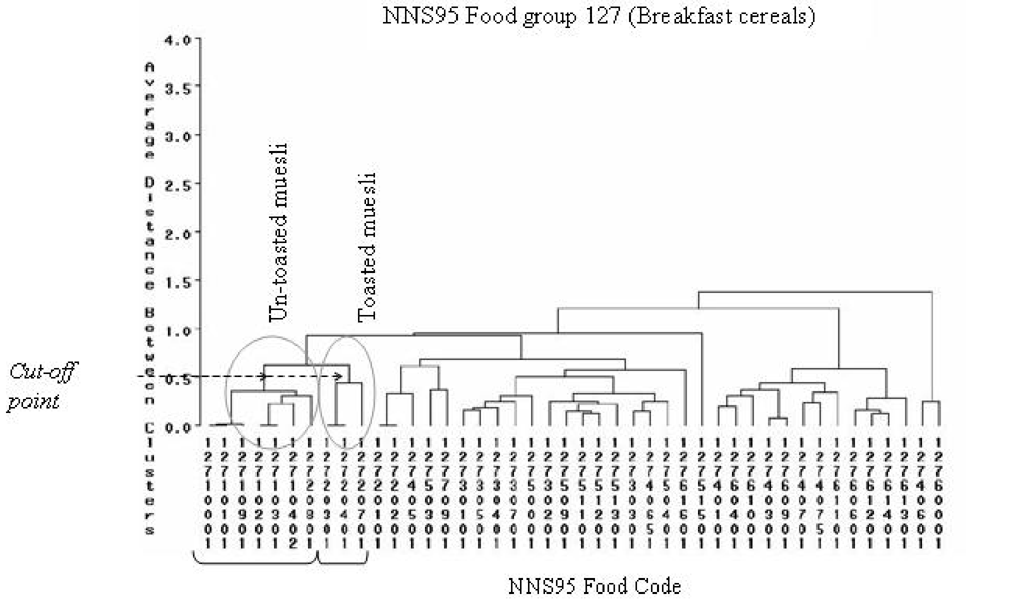
Table 5: Number of food groups in each level of the new food database

1st level food groups		*2nd level food groups*	*3rd level food groups*
Alcoholic drinks	1	4	15
Bakery products	2	9	27
Biscuits & crackers	3	2	12
Bread	4	5	25
Cereal	5	2	15
Convenience & takeaway foods	6	11	33
Dairy	7	6	31
Dishes with meat, chicken or fish	8	7	49
Eggs & egg dishes	9	1	6
Fruit & fruit dishes	10	3	13
Meal replacements & supplements	11	3	4
Meat, chicken & fish (not in a dish)	12	16	41

Non-alcoholic drinks	13	9	34
Rice & pasta dishes	14	2	21
Salad	15	1	14
Savoury snack foods	16	8	13
Soups	17	3	15
Sweet snack foods	18	4	12
Vegetables and vegetable dishes	19	7	42
Total number of food groups	**19a**	**103 b**	**422 c**

National Nutrition Survey (NNS95) hierarchy contained a 21 b 106 c 370 food groups

Figure 1: Cluster analysis for NNS95 group 127 (breakfast cereals) using average linkage clustering technique (Burden et al.). Figure shows cut-off points for grouping of food items and areas in which professional judgement was required for food group creation.



NNS95 Food Code	*NNS95 Food Group name*	*'New' food group*
12710001	Muesli, untoasted, < 20% total sugars, unfortified	Untoasted muesli
12710101	Muesli, untoasted, < 20% total sugars, fortified	Untoasted muesli
12710201	Muesli, untoasted, > 20% total sugars, unfortified	Untoasted muesli
12710301	Muesli, untoasted, > 20% total sugars, fortified	Untoasted muesli
12710402	Muesli, untoasted, homemade with dried fruit	Untoasted muesli
12710901	Muesli, untoasted, commercial, NFS	Untoasted muesli
12720101	Muesli, toasted, fat < 12%, fortified	Toasted muesli *
12720201	Muesli, toasted, fat <12%, unfortified	Toasted muesli *
12720301	Muesli, toasted, fat >12%, fortified	Toasted muesli
12720401	Muesli, toasted, fat >12%, unfortified	Toasted muesli
12720701	Muesli, toasted, NFS	Toasted muesli

* Professional judgment required. Abbreviations: NNS95 – National Nutrition Survey, NFS – No form specified.